A blue robot with glowing orange eyes and a white tablet is the background of the slide. The robot is holding the tablet with its right hand. The title 'Intelligent Agents Codebase' is written in yellow text on the tablet. The background is a solid blue color with some white lines.

Intelligent Agents Codebase

Group 3: Presentation



Contents

- Introduction
- Design
- Development Environment
- Code Quality
 - Formatting, Linting and Reviews
 - Testing, Coverage and Report
 - Metrics and Repository
 - Commenting and Documenting
- Codebase References
- BOT Running
- Explaining the ini configuration file
- Process & Code Flow
- Conclusion



Introduction

This presentation introduces with the coding approach and implementations for the “digital forensics” prototype design that is responsible to

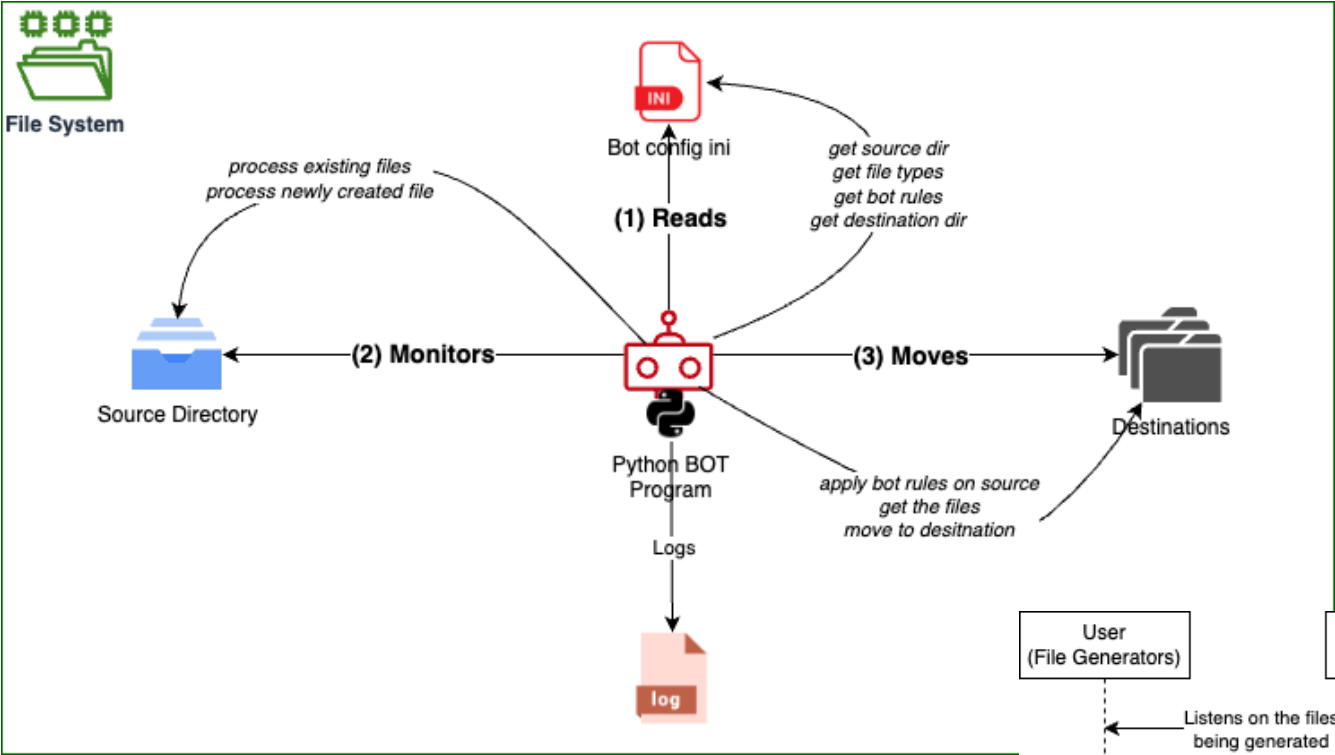
- monitor,
- identify,
- find
- and move files.

Based on the “iterative pattern”, the design proposed for the development of the BOT has undergone further team reviews to ensure that the BOT is self-sufficient and works using a configuration that becomes the data inputs for the BOT to process accordingly.

With regular team reviews of the approach and an understanding of the complexities, the development used a brief TDD (Test Driven Development) approach for the first rollout of the prototype. The initial release of the prototype (MVP1) was outlined to fulfil requirements that can be delivered on time as a baseline and enabler for future enhancements.

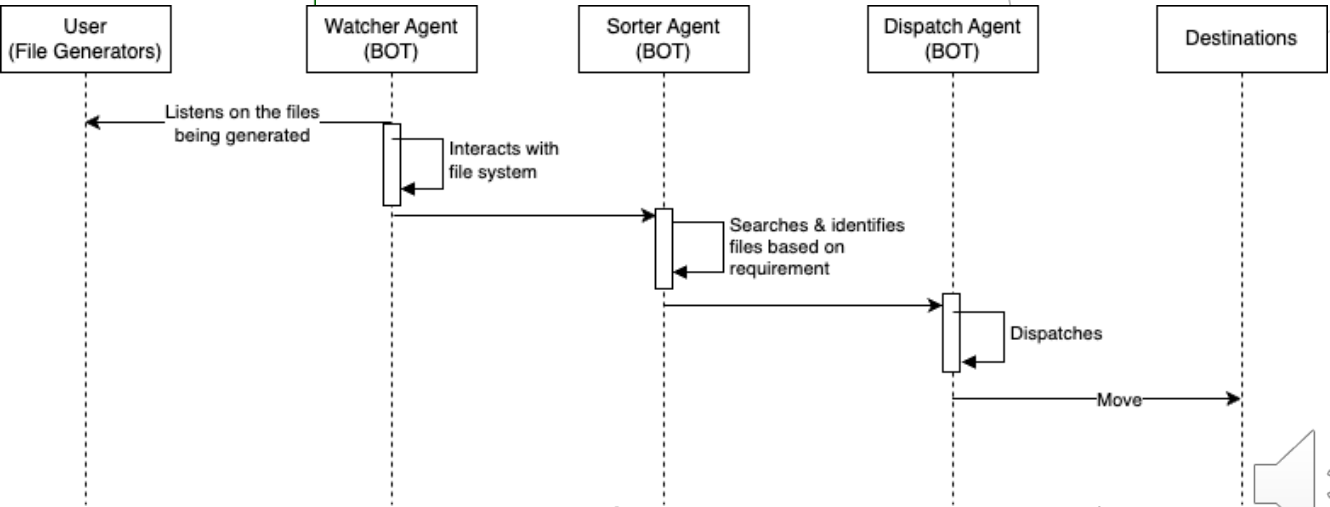


Design



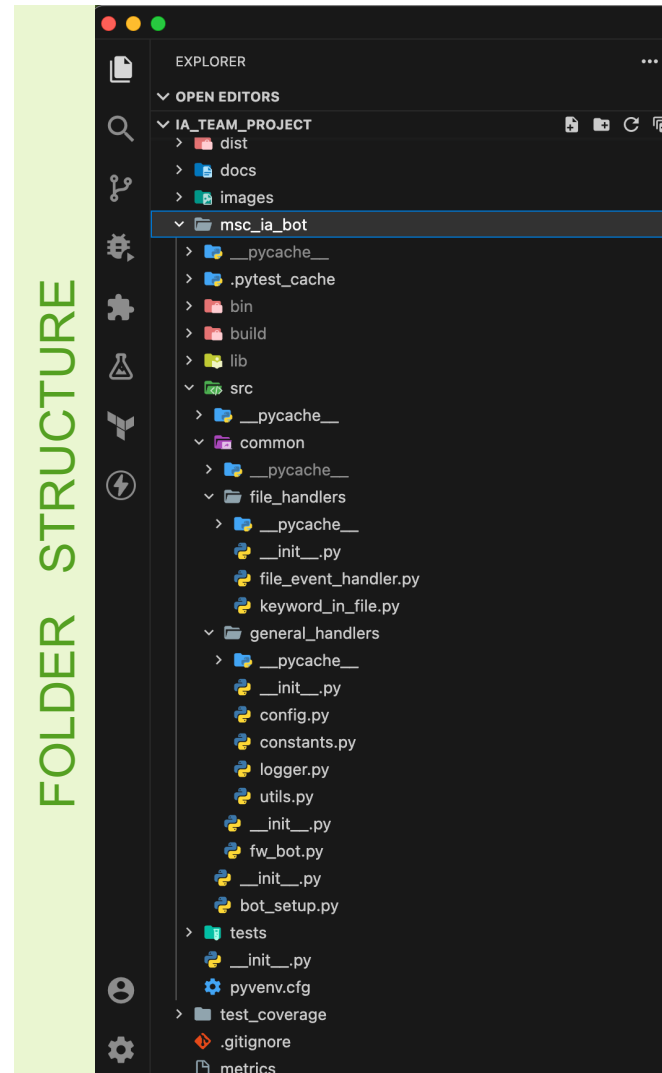
Three main functions:

- (1) READ
- (2) MONITOR
- (3) MOVE



Development Environment

- Virtual Environment
 - python 3.10.11
 - pip 23.1.2
 - pipdeptree 2.9.5
 - Code quality
 - pytest 7.4.0 (unit testing)
 - radon 6.0.1 (code metrics)
 - Sphinx 7.0.1 (code comments and documentation)
 - coverage 7.2.7 (code coverage)
 - Dependencies
 - watchdog 3.0.0
 - colorama 0.4.6
 - openpyxl 3.1.2
 - PyPDF2 3.0.1
 - Paramiko 3.2.0
 - Zipp 3.16.0
 - python-docx 0.8.11
 - urllib3 2.0.3
 - lxml 4.9.3
 - et-xmlfile 1.1.0
- Code Repository
 - Github
(https://github.com/biswassandip/IA_Team_Project)
- IDE
 - Visual Studio Code (vscode)
 - Extensions
 - autopep8
 - Pylance
 - Mypy Type Checker
 - Pylint

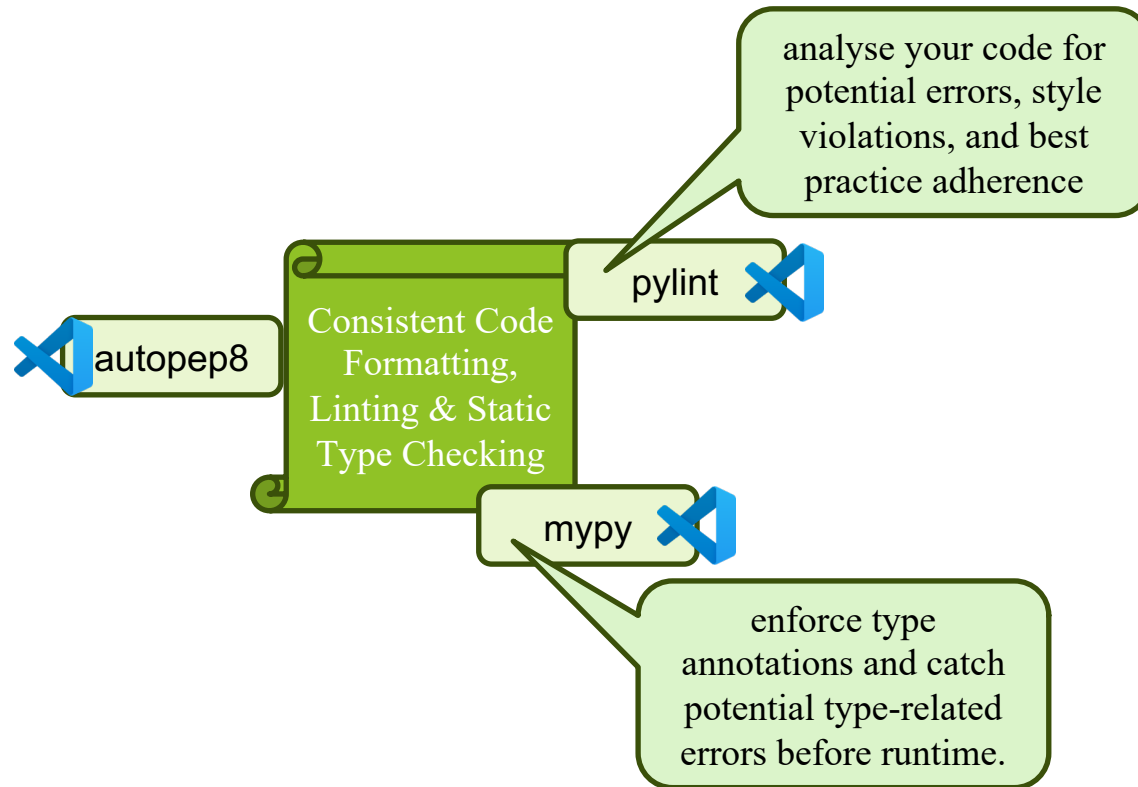


Setting up the environment

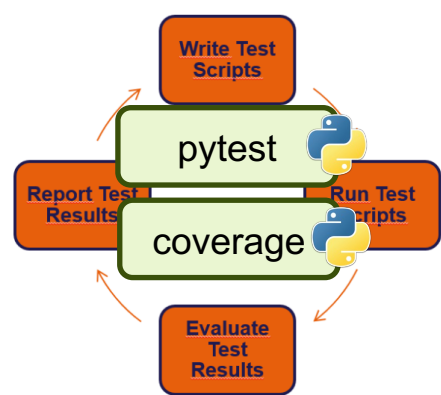
1. Create Git repository in Github
2. Clone the repository in development environment
3. Create the virtual environment (ex; msc_ia_bot)
4. Activate the virtual environment
5. Install dependencies and code quality libraries
6. Open the folder msc_ia_bot in vscode
7. Add vscode extensions
8. Select the interpreter in vscode
9. Create required folders
10. Happy coding



Code Quality (formatting, linting and reviews)



Code Quality (testing, coverage and report)



Refer:
https://htmlpreview.github.io/?https://github.com/biswassandip/IA_Team_Project/blob/main/test_coverage/index.html

Refer:
https://htmlpreview.github.io/?https://github.com/biswassandip/IA_Team_Project/blob/main/msc_ia_bot/tests/report.html

Coverage report: 79%
coverage.py v7.2.7, created at 2023-07-14 20:11 +0100

Module	statements	missing	excluded	coverage
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/__init__.py	0	0	0	100%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/__init__.py	0	0	0	100%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/bot_setup.py	101	48	0	52%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/common/__init__.py	0	0	0	100%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/common/general_handlers/__init__.py	0	0	0	100%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/common/general_handlers/file_event_handler.py	94	23	0	76%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/common/general_handlers/keyword_in_file.py	108	24	0	78%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/common/general_handlers/keyword_in_excel.py	35	3	0	91%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/common/general_handlers/keyword_in_pdf.py	0	0	0	100%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/common/general_handlers/keyword_in_zip.py	0	0	0	100%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/common/general_handlers/config.py	61	8	0	87%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/common/general_handlers/constants.py	29	0	0	100%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/common/general_handlers/logger.py	24	2	0	92%
/Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/src/common/general_handlers/validators.py	59	14	0	76%
__init__.py	0	0	0	100%
test_bot_processes.py	23	0	0	100%
test_bot_setup.py	27	4	0	85%
test_keyword_in_files.py	43	0	0	100%
Total	604	126	0	79%

coverage.py v7.2.7, created at 2023-07-14 20:11 +0100

Testing & coverage

report.html

Report generated on 15-Jul-2023 at 07:12:29 by `pytest-html`

Summary

13 tests ran in 10.44 seconds.

(Un)check the boxes to filter the results.

☒ 13 passed, ☒ 0 skipped, ☒ 0 failed, ☒ 0 errors, ☒ 0 unexpected passes

Results

Show all details / Hide all details

Result	Test
Passed (show details)	test_keyword_in_files.py::TestKeywordInFile::test_search_keyword_in_text_file
Passed (show details)	test_keyword_in_files.py::TestKeywordInFile::test_search_keyword_in_docx_file
Passed (show details)	test_keyword_in_files.py::TestKeywordInFile::test_search_keyword_in_excel_file
Passed (show details)	test_keyword_in_files.py::TestKeywordInFile::test_search_keyword_in_xml_file
Passed (show details)	test_keyword_in_files.py::TestKeywordInFile::test_search_keyword_in_pdf_file
Passed (show details)	test_keyword_in_files.py::TestKeywordInFile::test_search_keyword_in_zip_file
Passed (show details)	test_bot_processes.py::test_create_config
Passed (show details)	test_bot_processes.py::test_process_start
Passed (show details)	test_bot_processes.py::test_process_stop
Passed (show details)	test_bot_processes.py::test_process_existing_files
Passed (show details)	test_bot_setup.py::test_bot_config_setup
Passed (show details)	test_bot_setup.py::test_start_process
Passed (show details)	test_bot_setup.py::test_stop_process

Test Run Report

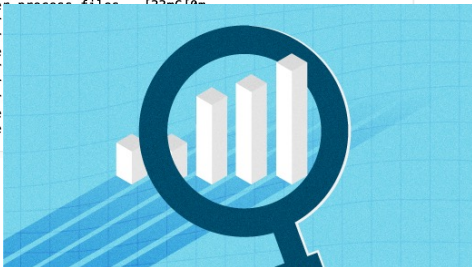
Code Quality (metrics and code repository)

Github:
https://github.com/biswassandip/IA_Team_Project/blob/main/msc_ia_bot/metrics

```
src/bot_setup.py
[1m[35mF 0m216:0 run - [32mB[0m
[1m[35mF 0m41:0 process_menu - [32mA[0m
[1m[35mF 0m108:0 process - [32mA[0m
[1m[35mF 0m78:0 bot_config_setup - [32mA[0m
[1m[35mF 0m159:0 start_process - [32mA[0m
[1m[35mF 0m187:0 stop_process - [32mA[0m
[1m[35mF 0m24:0 display_setup_menu - [32mA[0m
src/common/fw_bot.py
[1m[35mF 0m13:0 monitor_files - [32mB[0m
src/common/general_handlers/config.py
[1m[36mC 0m32:0 Config - [32mA[0m
[1m[37mM 0m68:4 Config.create_config - [32mA[0m
[1m[37mM 0m150:4 Config.generate_ssh_key - [32mA[0m
[1m[37mM 0m49:4 Config.__init__ - [32mA[0m
src/common/general_handlers/logger.py
[1m[36mC 0m19:0 Logger - [32mA[0m
[1m[37mM 0m22:4 Logger.__init__ - [32mA[0m
[1m[37mM 0m28:4 Logger.init_logger - [32mA[0m
[1m[37mM 0m50:4 Logger.info - [32mA[0m
[1m[37mM 0m62:4 Logger.warning - [32mA[0m
[1m[37mM 0m74:4 Logger.error - [32mA[0m
src/common/general_handlers/utls.py
[1m[37mM 0m22:4 Utlis.__print_it - [32mA[0m
[1m[36mC 0m18:0 Utlis - [32mA[0m
[1m[37mM 0m172:4 Utlis.create_d - [32mA[0m
[1m[37mM 0m120:4 Utlis.validate - [32mA[0m
[1m[37mM 0m66:4 Utlis.custom_in - [32mA[0m
[1m[37mM 0m80:4 Utlis.custom_print - [32mA[0m
[1m[37mM 0m94:4 Utlis.information - [32mA[0m
[1m[37mM 0m107:4 Utlis.error_message - [32mA[0m
[1m[37mM 0m191:4 Utlis.full_path - [32mA[0m
[1m[37mM 0m206:4 Utlis.get_file_name - [32mA[0m
src/common/file_handlers/keyword_in_file.py
[1m[37mM 0m24:4 KeywordInFile.search_keyword_in_files - [32mB[0m
[1m[37mM 0m116:4 KeywordInFile.search_keyword_in_excel_file - [32mB[0m
[1m[36mC 0m18:0 KeywordInFile - [32mB[0m
[1m[37mM 0m205:4 KeywordInFile.search_keyword_in_zip_file - [32mB[0m
[1m[37mM 0m88:4 KeywordInFile.search_keyword_in_docx_file - [32mA[0m
[1m[37mM 0m177:4 KeywordInFile.search_keyword_in_pdf_file - [32mA[0m
[1m[37mM 0m61:4 KeywordInFile.search_keyword_in_text_file - [32mA[0m
[1m[37mM 0m150:4 KeywordInFile.search_keyword_in_xml_file - [32mA[0m
[1m[37mM 0m21:4 KeywordInFile.__init__ - [32mA[0m
src/common/file_handlers/file_event_handler.py
[1m[37mM 0m127:4 FileMoveEventHandler - [32mB[0m
[1m[36mC 0m19:0 FileMoveEventHandler - [32mA[0m
[1m[37mM 0m68:4 FileMoveEventHandler - [32mA[0m
[1m[37mM 0m107:4 FileMoveEventHandler - [32mA[0m
[1m[37mM 0m31:4 FileMoveEventHandler - [32mA[0m
[1m[37mM 0m49:4 FileMoveEventHandler - [32mA[0m
[1m[37mM 0m92:4 FileMoveEventHandler - [32mA[0m
[1m[37mM 0m202:4 FileMoveEventHandler - [32mA[0m
[1m[37mM 0m186:4 FileMoveEventHandler - [32mA[0m
```

Metrics
(Cyclomatic Complexity)

radon



github.com/biswassandip/IA_Team_Project

Cloud vs... Home - Digital Clo... AWS vs. Azure vs... AWS to Azure ser... Map Amazon AWS... Amazon Kinesis D...

Pull requests Actions Projects Wiki Security Insights Settings

IA_Team_Project Public Pin Unwatch

Repository:
https://github.com/biswassandip/IA_Team_Project 66 commits

images	process	2 days ago
msc_ia_bot	installation	2 days ago
.DS_Store	process	2 days ago
.gitignore	dist	2 days ago
readme.md	readme	2 days ago
requirements.txt	installation	2 days ago

readme.md

Digital File Monitoring BOT

python >=3.10 dist download release v1.0.0 docs sphinx htmlcov coverage test case pytest metrics radon

The BOT is responsible to Monitor, Retrieve and Process the Source Directory based on a Configuration ini file as the input. This input is required with proper definitions about the keywords and/or file-types so that the data can be used to move the files to respective directories so that they can be given for further analysis.



Code Quality (code comments and documenting)

```
bot_setup.py x
IA_Team_Project > msc_ia_bot > src > bot_setup.py > ...
1 """
2 **Module:** bot_setup.py
3
4 This module is the start module that has the responsibilities
5
6 It provides the user with a menu that has options to choose from
7
8 ...* Setup requirements to create an ini file that the bot will use
9 ...* Start the bot.
10 ...* Stop the bot.
11 ...* Quit the setup.
12
13 Based on the chosen options, this module will decide on creating the
14 The **User** is expected to review the ini file and update it as required
15 to start the process.
16 """
17
18 from common.general_handlers.utils import Utils
19 from common.general_handlers.config import Config
20 import configparser
21 import common.fw_bot as fw_bot
22 import sys
23
24 def display_setup_menu():
25     """
26     This method is responsible to display the menu options.
27     """
28
29     Utils.custom_print(f"=====
30     Utils.custom_print(f"BOT SETUP MENU")
31     Utils.custom_print(f"=====
```

sphinx



File Sorter

BOT

Navigation

Contents:

src

Quick search

 Go

Python Module Index

/TeamProject/IA_Team_Project/msc_ia_bot/docs/_build/html/src.html#module-src.bot_setup

s

src

[src.bot_setup](#)

[src.common](#)

[src.common.fil](#)

[src.common.fil](#)

[src.common.fil](#)

[src.common.fw](#)

[src.common.gen](#)

[src.common.gen](#)

[src.common.gen](#)

[src.common.gen](#)

[src.common.gen](#)

Github download -

https://github.com/biswassandip/IA_Team_Project/tree/main/msc_ia_bot/docs/_build/html
folder

src.bot_setup module

Module: bot_setup.py

This module is the start module that has the responsibilities to fulfill the setup of the bot.

It provides the user with a menu that has options to choose from:

- Setup requirements to create an ini file that the bot will use.
- Start the bot.
- Stop the bot.
- Quit the setup.

Based on the chosen options, this module will decide on creating the required ini file (**bot_config.ini**). The **User** is expected to review the ini file and update it as required before rerunning the bot_setup.py to start the process.

src.bot_setup.**bot_config_setup**() [\[source\]](#)

This method will trigger the creation of the actual config ini file.

Returns: boolean

src.bot_setup.**display_setup_menu**() [\[source\]](#)

This method is responsible to display the menu options.

src.bot_setup.**process**(b_start, config_file_path, sleep_time=5, b_execute_once=False) [\[source\]](#)

This is function processes based on chosen option 2 to start and 3 to stop.



Codebase references

- Codebase - The entire code has been checked into Github. Click [here](#)
- Test Cases - The test cases from this project. Click [here](#)
- Test Run Report - This provides the execution report and results. Click [here](#)
- Test Coverage Report - This shows how much the code has been covered after running tests. Click [here](#)
- Documentation - The entire code documentation and structure can be found here. Download the html folder. Click [here](#)



BOT Running



Explaining the ini configuration file

GENERAL:

Can be updated with the directory to be monitored (source_dir) and where the logs to be created

SEARCH_IN_FILE_TYPES:

List of comma separated file types that will be used to search for the keyword

BOT_RULES:

A "criteria" or rule which is a comma separated string of the following: Keyword, file type, sftp host, sftp port, destination directory.

The SFTP Logic is an enhancement requirement to current BOT that will be implemented in next releases after reviewing the requirements more in details.

The PROCESSES are for multiprocessing enhancement for later

FLAGS -

stop_flag = True would mean that the process will be stopped

IA_BOT >  bot.ini

[GENERAL]

```
ini_file_path = /Users/gini/TeamProject/IA_Team_Project/dist/IA_BOT/bot.ini
source_dir = /Users/gini/TeamProject/IA_Team_Project/msc_ia_bot/tests/TEST_PAT
log_file = /Users/gini/TeamProject/IA_Team_Project/dist/IA_BOT/log/bot_log.log
rotate_logs = True
rotation_size = 1000000
```

[SEARCH_IN_FILE_TYPES]

```
include = .txt, .doc, .docx, .xls, .xlsx, .xml, .pdf, .png, .zip, .html, .css
```

[BOT_RULES]

```
criteria1 = example, .pdf, , , /Users/gini/TeamProject/IA_Team_Project/msc_ia
criteria2 = example, .txt, , , /Users/gini/TeamProject/IA_Team_Project/msc_ia
criteria3 = example, .xlsx, , , /Users/gini/TeamProject/IA_Team_Project/msc_ia
criteria4 = example, .docx, , , /Users/gini/TeamProject/IA_Team_Project/msc_ia
```

[SFTP_KEYS]

```
pf_file = ./IA_BOT/keys/private_key.pem
```

[PROCESSES]

```
num_processes = 4
num_processes_scaling_factor = 0.8
min_processes = 2
```

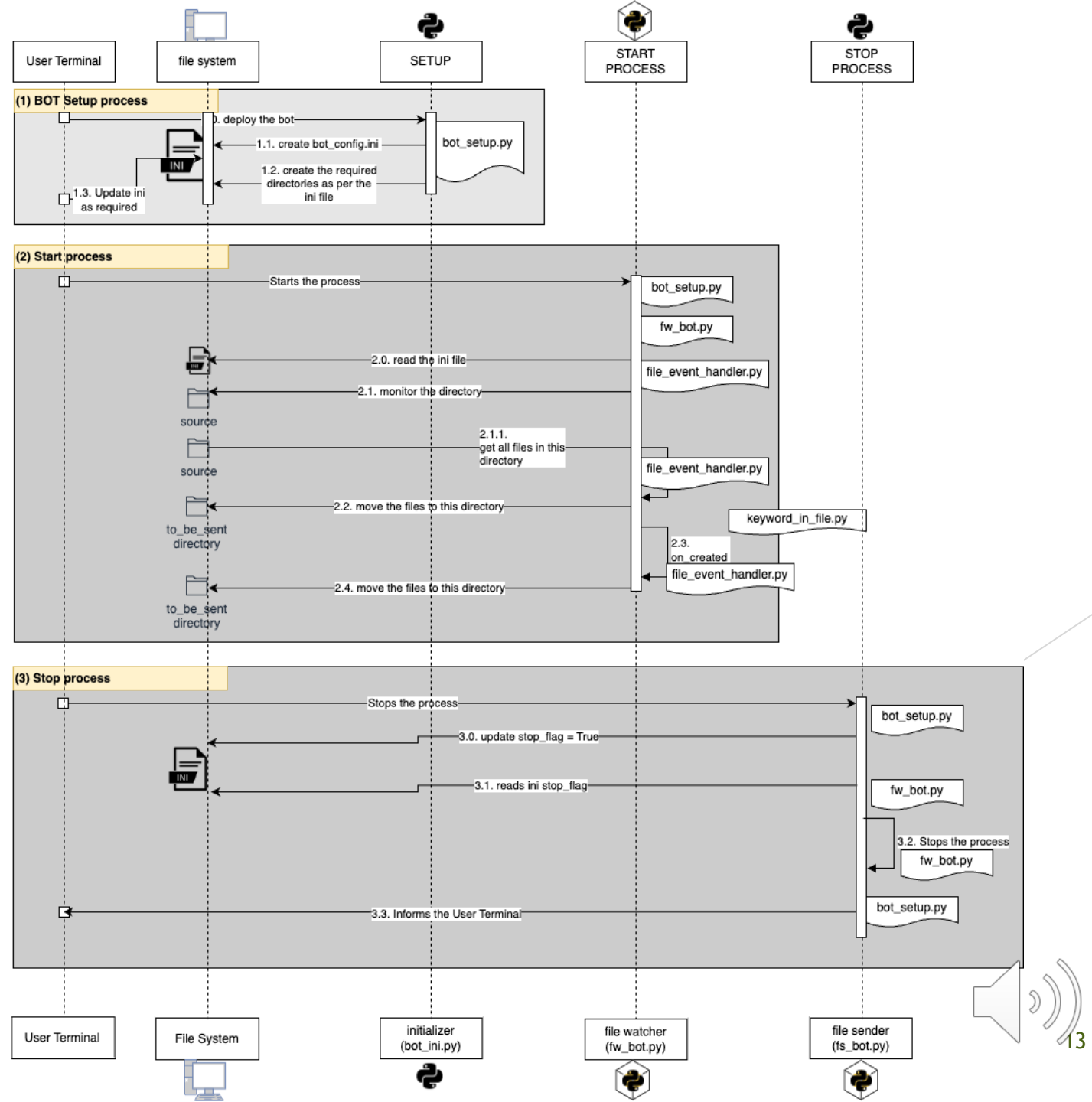
[FLAGS]

```
stop_flag = True
```

Process & Code Flow

A brief sequence is provided of when a python program is executed during a process:

- **SETUP Process**
 - Responsible for creation of the configuration ini file
 - The User is required to review this file and update as required (refer the configuration ini slide for details)
- **START Process**
 - Once the BOT has been started, it reads the ini file and performs the activities as per the BOT rules provided in the ini file
 - It is running until stopped
- **STOP Process**
 - The running BOT can be stopped by either updating the stop_flag to True manually in the ini file or invoking the setup



Conclusion

The initial MVP for the Digital File Handling BOT has been built with a baseline structure and framework for future enhancements like adding SFTP and REST features for destinations. This would call for adding respective features in the configuration.

This BOT is self-sufficient and can manage based on the data inputs, thus reducing the time to manually move files based on keyword searches within the file or file types. A focused approach to refine and define the BOT rules in the configuration file would enable faster and easier management of the files.

The BOT results are supported by the coverage and test reports that can be viewed online. Moreover, the entire documentation for the packages and modules has also been built and provided through the repository.

Industry standards have been used to make sure that the initial release of the BOT codebase is maintainable and can be enhanced with better measures.

